



TITLE:

# A Study on Resolution and Retrieval of Implicit Entity References in Microblogs( Abstract\_要旨 )

AUTHOR(S):

Lu, Jun-Li

---

CITATION:

Lu, Jun-Li. A Study on Resolution and Retrieval of Implicit Entity References in Microblogs.  
京都大学, 2020, 博士(情報学)

ISSUE DATE:

2020-03-23

URL:

<https://doi.org/10.14989/doctor.k22580>

RIGHT:

様式VI

## 博士学位論文調査報告書

論文題目 A Study on Resolution and Retrieval of Implicit Entity  
References in Microblogs  
(マイクロブログにおける暗黙的な実体参照の解決および検索に関する研究)

申請者氏名 Jun-Li LU

最終学歴 平成25年9月  
京都大学大学院情報学研究科 社会情報学専攻修士課程 修了  
  
平成29年3月  
京都大学大学院情報学研究科 社会情報学専攻博士後期課程  
研究指導認定退学

学識確認 平成 年 月 日 (論文博士のみ)

論文調査委員 京都大学大学院情報学研究科  
(調査委員長) 教授 吉川 正俊

論文調査委員 京都大学大学院情報学研究科  
教授 黒橋 禎夫

論文調査委員 京都大学大学院情報学研究科  
教授 田島 敬史

論文調査委員 京都大学大学院情報学研究科  
名誉教授 田中 克己

( 続紙 1 )

京都大学	博士（情報学）	氏名	Jun-Li Lu
論文題目	A Study on Resolution and Retrieval of Implicit Entity References in Microblogs （マイクロブログにおける暗黙的な実体参照の解決および検索に関する研究）		
<p>（論文内容の要旨）</p> <p>In the age of social media, people share their opinions about the real world by writing microblogs in a SNS platform (e.g. Twitter or Facebook) with references to entities (such as products, people) and events. Analyzing those references in microblogs can be useful for various kinds of applications, such as short text (microblogs) retrieval, targeted advertising and a low-cost way for public opinion survey. Conventional research work has already extensively studied <i>explicit</i> entity references in microblogs, which are based on literal text (e.g. names or keywords) of entities. However, <i>implicit</i> reference cases, which are not based on literal text of entities, have been increasing in SNS platforms and not been fully studied yet. This thesis studies the resolution and retrieval of implicit entity references in microblogs from multiple aspects.</p> <p>Chapter 1 outlines the thesis, including the research background of the resolution and retrieval of implicit entity references, motivation of this research, tasks involved in this thesis, and an overview of the thesis.</p> <p>Chapter 2 discusses and overviews previous studies related to our research topics presented in this thesis.</p> <p>Chapter 3 addresses the problem of identifying multiple indirectly or directly referred entities in a microblog. The author especially tackles indirect reference cases by developing novel features, such as topic-based semantics. To accurately model dependencies among entities and achieve high computational efficiency, we model dependencies by a conditional random field model with adaptive dependency.</p> <p>Chapter 4 addresses the problem of identifying implicitly or explicitly referred events in microblogs. To this end, reference-type-based event’s contents generation is utilized. To generate contents for implicitly referred events, similar links among microblogs and events are mined by clustering them respectively. The author also predicts the reference type for accurately identifying implicitly or explicitly referred events.</p> <p>Chapter 5 proposes a novel approach for event identification. Since microblogs and events are different in their content, a deep transfer learning framework is introduced that dually transforms microblogs and events respectively such that the transformed microblogs and events are easier to be matched.</p>			

Chapter 6 addresses the problem of searching for microblogs implicitly or explicitly referring to events, through a search API. To this end, a new search process is introduced that takes dynamic query strategies optimized by search experiences, with the following components. (a) Exploratory and exploitative query strategies are introduced to search for microblogs implicitly or explicitly referring to events. (b) To accurately monitor the search process, a robust reinforcement learning framework is proposed that predicts long-term returns of search process by using current and historical search experiences. (c) To increase the applicability of this research, online query strategies are optimized by estimating the relevance of microblogs using the relevance estimator proposed in Chapter 5.

Chapter 7 summarizes the thesis and addresses several directions to be explored in future work.

(論文審査の結果の要旨)

近年、現実世界の物事に対する意見の多くがTwitterやFacebookなどのソーシャルネットワークサービスにて共有されるようになった。特にTwitterなどへの投稿はマイクロブログとよばれ、マイクロブログを検索・解析することによって、ターゲティング広告や低コストの意見調査など様々なアプリケーションを実現できる。マイクロブログ中では、人物や商品などのエンティティ（実体）やイベント（事象）が参照されており、マイクロブログの内容を十分に理解するためには、ある言語表現からどのエンティティやイベントが参照されているのかを明らかにする必要がある。この問題に対して、エンティティの名前による参照、すなわち、明示的な参照については多くの研究がなされてきている。しかしながら、エンティティやイベントの名前によらない参照、すなわち、暗黙的な参照については、これまで広く研究が行われてこなかった。本論文では、この暗黙的な参照の解消および検索について4つの問題に取り組み、各問題について以下の成果を上げている。

1. 適応的な依存関係を考慮した条件付き確率場によるエンティティ同定

マイクロブログに含まれる暗黙的なエンティティ参照を解消し参照エンティティを同定する問題に取り組んだ。この暗黙的な参照解消問題を解決しエンティティを同定するために、適応的な依存関係を考慮した条件付き確率場を用いた手法を提案した。実験では、明示的および暗黙的なエンティティの参照を含むデータセットを構築し、適応的な依存関係を考慮した提案手法が効果的であることを示した。

2. 明示的および暗黙的参照からのイベント同定

マイクロブログに含まれる明示的および暗黙的なイベント参照を解消する問題に取り組んだ。特に、暗黙的なイベント参照については、参照元がイベント中の単語を含まないことが課題であった。この問題に対して、あるマイクロブログの参照先のイベントに類似するような仮想的なイベントを生成することによって、語彙の差異を埋める方法を提案し、その有効性を示した。

3. 双方向変換を行う深層学習によるイベント同定

明示的および暗黙的なイベントの参照を解消する問題について、双方向変換を行う深層学習モデルを提案した。このモデルでは、マイクロブログの内容とイベントの内容をそれぞれ別のモデルによって、同一の空間に写像し、その空間上での類似度によって、あるマイクロブログがどのイベントを参照しているかを判定することができる。

4. 動的クエリ戦略に基づいたイベント参照マイクロブログの検索

あるイベントを参照するマイクロブログをマイクロブログサービスの検索APIを通して効率的に収集するために、参照されるイベントや得られた検索結果に応じて動的に変わるクエリ戦略によって検索を行う方法を提案した。実験では、動的なクエリ戦略が静的なクエリ戦略よりも優れ、効率的なマイクロブログ収集が可能であることを示した。

以上、本論文では、人、組織、商品などの一般的なエンティティ（実体）の暗黙的参照の解消問題、動作やできごとを表すイベント（事象）の暗黙的参照の解消問題、および、イベントへの参照を含むマイクロブログを検索する問題に取り組み、成果を上げたものである。この研究成果は、マイクロブログ等に特徴的な短文・会話文の検索・解析に資するもので、学術上、および、實際上、寄与するところが少なくない。よって、本論文は博士（情報学）の学位論文として価値あるものと認める。また、令和2年2月27日に実施した論文内容とそれに関連した試問の結果、合格と認めた。

Webでの即日公開を希望しない場合は、以下に公開可能とする日付を記入すること。  
要旨公開可能日： 年 月 日以降